Statistik

in Sozialer Arbeit und Pflege

Von
Prof. Dr. Rüdiger Ostermann
und
Prof. Dr. Karin Wolf-Ostermann

3., überarbeitete Auflage

Oldenbourg Verlag München

Bibliografische Information Der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

1. Nachdruck 2012

Rosenheimer Straße 145, D-81671 München
Telefon: (089) 45051-0
www.oldenbourg-verlag.de

Gedruckt auf säure- und chlorfreiem Papier
Gesamtherstellung: Books on Demand GmbH, Norderstedt

ISBN 3-486-57763-8
ISBN 978-3-486-57763-1
eISBN 978-3-486-59947-3

Inhalt

Vorwort

Wir haben uns sehr gefreut, dass unsere 2. Auflage ebenfalls auf ein so großes Interesse gestoßen ist, dass wir nun mit dem Oldenbourg-Verlag vereinbaren konnten, eine dritte, überarbeitete und korrigierte Auflage herauszubringen. So haben wir insbesondere die Praxisbeispiele nicht mehr nur aus dem Bereich der Sozialen Arbeit, sondern auch aus dem Bereich der Pflege gewählt, da beide Autoren seit einiger Zeit auch in diesem Bereich tätig sind.

Unseren bisherigen Anhang haben wir überarbeitet und zu einem „richtigen" Kapitel aufgewertet. Dieses ist nun allgemein dem empirischen Arbeiten gewidmet und als zweites Kapitel in das vorliegende Buch integriert. Zusätzlich haben wir dem Wunsch einer Buchbesprechung zu unserer 2. Auflage entsprochen und den Bereich der Stichproben erheblich ausgeweitet. Ein weiteres Anliegen für uns war es, explizit auf die Gestaltung von Tabellen einzugehen, da wir den Eindruck gewonnen haben, dass dieses Thema in vielen Lehrbüchern vernachlässigt wird. Das schon vorhandene Indexregister wurde so weit vervollständigt, dass man unser Buch auch als Nachschlagewerk verwenden kann.

Bedanken möchten wir uns bei den vielen Studierenden unserer Lehrveranstaltungen, die uns seit dem Erscheinen der 1. Auflage auf Rechtschreib- und Rechenfehler hingewiesen haben. Unser Dank gilt auch – wie schon bei der ersten und zweiten Auflage – dem Oldenbourg-Verlag und insbesondere Herrn Diplom-Volkswirt M. Weigert, dass sie es weiterhin „wagen, mit unserem Werk auf den Büchermarkt zu kommen".

Aus dem Vorwort zur 2. Auflage

Es sei noch einmal darauf hingewiesen, dass wir uns bemüht haben, etwaige mathematische Herleitungen in einfachen Schritten zu erklären, so dass sie ggf. sogar übersprungen werden können, ohne dass das weitere Lesen und Verständnis dadurch behindert wird. Wir wollen es ein wenig wie *Gilmore (1996)* halten, der in seinem Vorwort schreibt: „Damit haben wir auch schon den Grad mathematischer Kompliziertheit in diesem Buch gezeigt. Man könnte mit einem mathematischen Ausdruck noch sehr viel genauer sein. Ein Ausdruck wie dieser zum Beispiel

$$\frac{\partial}{\partial t}\iiint_{\Delta_i} P(r,t)d^3r = \iint_{S\Delta} \Im_P \cdot \hat{n}dS$$

enthält sehr viel Information für diejenigen, denen eine solche Schreibweise vertraut ist. Aber für uns ist das nichts. Die bloße Anwesenheit von Mathematik ist vor allen Dingen nicht ansteckend. Sie brauchen die Gleichungen nicht zu lesen, wenn Sie nicht wollen. Die Erklärungen im Text sind die gleichen, ob mit oder ohne Mathematik. Damit Sie sich, wenn Sie wollen, auf Zehenspitzen zurückschleichen können, ohne die mathematischen Ungeheuer aufzuwecken, ist vor jeder Formel ein Warnzeichen angebracht:

Achtung Mathematik!

Es zeigt an, dass eine mathematische Gleichung kommt."

Ganz so weit wollen wir nicht gehen, und verzichten deshalb auf die Einführung dieses Warnzeichens. Nichtsdestotrotz wünschen wir unsern Leserinnen und Lesern viel Freude bei der Lektüre unseres Buches.

Aus dem Vorwort zur 1. Auflage

In den letzten Jahren sind zahlreiche Lehrbücher zur Einführung in die Statistik erschienen, die sich jedoch nach unserer Einschätzung nicht allzu sehr an einen Leser- und Nutzerkreis aus dem Bereich der Sozialpädagogik oder Sozialarbeit wenden. Begründet auf ein von uns erstelltes Vorlesungsmanuskript zu einer Lehrveranstaltung „Einführung in die Statistik", die wir beide seit einer Reihe von Semestern an der Universität-GH Siegen im Fach „Außerschulisches Erziehungs- und Sozialwesen" geben, haben wir uns entschlossen, dieses Vorlesungsmaterial gründlich zu überarbeiten und in eine druckfertige Form zu bringen.

Wir haben uns bemüht, das „statistische Denken" in einfacher und verständlicher Form dazulegen. Für das notwendige Verständnis beim Durcharbeiten dieses Buches setzen wir kein hohes mathematisches Vorwissen voraus. Etwaige mathematische Herleitungen werden in einfachen Schritten erklärt oder können ggf. sogar übersprungen werden, ohne dass das weitere Lesen und Verständnis dadurch behindert wird. Des Weiteren haben wir zu jedem eingeführten Begriff ein oder mehrere Beispiele ausführlich ausgearbeitet. Jedes Kapitel wird von einem Aufgabenblock abgeschlossen. Die Lösungen dazu findet man im Anhang. Zur Abrundung haben wir noch einige historische Fußnoten eingearbeitet, damit man ein geschlossenes Bild vom Wissensgebiet „Statistik" erhält.

R. Ostermann & K. Wolf-Ostermann

1 Einleitung

1.1 Wie viel Statistik benötigt man in der Sozialen Arbeit bzw. in der Pflege?[1]

Hinter der Frage wie viel Statistik man in diesen Arbeitsfeldern benötigt steckt die folgende Überlegung: Inwieweit muss sich ein(e) Mitarbeiter(in) im Bereich der Sozialen Arbeit bzw. der Pflege mit den Methoden und Verfahren der Statistik für ihre/seine tägliche Arbeit auskennen?

In Zeiten der Geldknappheit wird bzgl. zahlreicher sozialer Einrichtungen und Dienstleistungen hinterfragt, inwieweit es sinnvoll ist, diese Einrichtung weiterhin bestehen zu lassen oder alternativ zu schließen. Des Weiteren stellt sich immer öfter die Frage, welches von mehreren beantragten Projekten bewilligt oder welches der schon laufenden Projekte weiter finanziert werden soll. Auch stellt sich die Frage, ob und wie die Qualität der geleisteten Arbeit gemessen und bewertet werden kann.

Bei derartigen Fragestellungen benötigen die Mitarbeiter(innen) nicht nur Argumente, die auf empirischem Zahlenmaterial beruhen, sondern auch angemessene Präsentationstechniken (Tabellen, Grafiken, etc.), um auf geeignete Art und Weise zu reagieren. Aber auch die Leitungskräfte der einzelnen Einrichtungen sowie die Personen, die die zugehörigen politischen Entscheidungen treffen, müssen bzw. sollten in der Lage sein, aufgrund der präsentierten Ergebnisse empirischer Studien Entscheidungen zu treffen.

Tatsächlich kann man jedoch nur dann Daten oder Ergebnisse interpretieren und Vor- und Nachteile einer statistischen Methode erkennen und beurteilen, wenn man – wenigstens in groben Zügen – den Aufbau der jeweiligen statistischen Methode verstanden hat. Trotz eines (nach der Lektüre des Buches hoffentlich) soliden statistischen Grundwissens, sollte man sich jedoch nicht scheuen, in den Fällen, in denen man erkennt, dass das eigene Wissen zur Behandlung des statistischen Problems nicht ausreicht, einen statistischen Experten zu Rate zu ziehen. Getreu dem Motto *„Wenn ich ernstlich krank bin, suche ich einen Arzt auf und*

[1] Diese Überschrift orientiert sich an dem Titel einer 1997 erschienenen gemeinsamen Arbeit der Autoren, die sich mit Fragen zur Statistikausbildung für Mitarbeiter(innen) im Bereich der Sozialen Arbeit bzw. der Soziale Dienste befasst.

kuriere mich nicht selbst!" ist es für eine(n) Sozialarbeiter(in) bzw. für eine(n) Pflegemitarbeiter(in) nicht diskriminierend den Statistiker aufzusuchen. Die Hinzuziehung eines statistischen Experten ist in vielen Fällen zwar mit Kosten verbunden, es ist jedoch dabei abzuwägen, ob dies im Endeffekt nicht doch preisgünstiger ist, als

- eine Mitarbeiterin oder einen Mitarbeiter über einen langen Zeitraum für die statistische Analyse frei zu stellen, wo sie/er an anderer Stelle besser eingesetzt werden könnte und/oder
- durch mangelnde statistische Kenntnisse eine falsche bzw. unzureichende Analyse durchzuführen, aus der dann falsche, d.h., kostenintensive Entscheidungen abgeleitet werden.

Die Suche nach einem geeigneten statistischen Experten kann sich unter Umständen schwierig gestalten, jedoch gibt es gerade im Umfeld von Universitäten und Fachhochschulen genügend Gelegenheit fündig zu werden. So existiert in vielen Fällen im Bereich der Rechenzentren oder im Bereich von statistischen Lehrstühlen eine statistische Beratungsstelle. Es besteht aber auch die Möglichkeit sich direkt an die Lehrenden einer Hochschule zu wenden, die für die Fachgebiete Statistik / Empirische Sozial- oder Pflegeforschung zuständig sind.

1.2 Was ist Statistik?

Wenn man in einer deutschen Enzyklopädie, z.B. im *Brockhaus*, unter dem Begriff Statistik nachschlägt, so findet man:

Statistik [18.Jahrh. aus ital.], im *materiellen* Sinn die geordnete Menge von Informationen in Form empirischer Zahlen (>Statistiken<); im *instrumentalen* Sinn (Statist. Methoden) der Inbegriff der Verfahren, nach denen empirische Zahlen gewonnen, dargestellt, verarbeitet, analysiert und für Schlußfolgerungen, Prognosen und Entscheidungen verwendet werden. Im ersteren Sinn hat sich die S. institutionalisiert (Statist. Institutionen). Im letzteren Sinn ist sie ein wissenschaftl. Fach, das an Universitäten gelehrt wird. Im beide Bedeutungen umfassenden Sinn kann die S. definiert werden als >Inbegriff theoretisch fundierter, empirischer, objektivierter Daten< (G. MENGES). Danach gehören zur S. die → Daten, die an theoret. Konzepten orientiert und aus der Realität nach bestimmten Meßvorschriften gewonnen sind, sowie die Verfahren, die zu solchen Daten führen und die solche Daten zum Gegenstand haben.

Der Begriff S. geht auf M. → Schmeitzel (1679 bis 1747) zurück, der in Jena und Halle Vorlesungen (Titel: >Collegium politico-statisticum<) hielt. Das Wort S. wurde von G. → Achenwall (1719-72) geprägt; er leitete es von dem italien. >statista< (Staatsmann) ab und verstand darunter das Wissen, das ein Staatsmann besitzen sollte. In dieser Bedeutung als >Praktische Staatskunde< hielt sich der Begriff ein Jahrhundert. Um die Mitte des 19.Jahrh. wandelte und teilte er sich zu seinen heute verwendeten Bedeutungen.

Etwas weiter hinten findet man dann unter dem Subbegriff *Geschichte*: Die ältesten Zeugnisse statist. Tätigkeit reichen ins dritte vorchristliche Jahrtsd. zurück, die erste bekanntgewordene statist. Ermittlung wurde um 3050 v. Chr. in Ägypten durchgeführt. Mehrfach fanden Zählungen in der röm.

Kaiserzeit und der Zeit des vorspan. Inkareichs statt; für das MA. liegen vergleichsweise wenig S. vor. In der Form systemat. Staatenbeschreibungen gab es S. zwischen dem 16. und 18.Jahrh., zunächst in Italien, später in Holland und Dtl. Statist. Erhebungen über den auswärtigen Handel wurden schon im 17.Jahrh. veranstaltet und später umfangreich ausgebaut, doch wurden deren Ergebnisse meist geheimgehalten. Die erste Schule der wissenschaftl. S. (>Universitätsstatistik<) wurde 1660 von Hermann → Conring an der Universität Helmstedt begründet, G. → Achenwall lehrte seit 1748 in Göttingen. Mit moderner S. hatte diese Staatenkunde jedoch wenig gemeinsam, da sie auf quantitative Aussagen weitgehend verzichtete. Zu Beginn des 19.Jahrh. zerfiel diese Richtung als Folge einerseits der Arbeitsteilung (Staatslehre, Geographie, Volkswirtschaftslehre) und andererseits der wachsenden Zahl von nationalen statist. Ämtern, die die Zahlensuche privater Forscher überflüssig machte. Die Disziplin überlebte durch die sich 1830-50 vollziehende Verbindung mit der >Politischen Arithmetik<, die in England schon in der Mitte des 17.Jahrh. von John → Graunt und William → Petty begründet worden war; ihr Vertreter in Dtl. war J.P. → Süßmilch. Im Gegensatz zur deskriptiv orientierten Universitäts-S. war die Politische Arithmetik auf der Suche nach Regelmäßigkeiten im Wirtschafts- und Sozialleben. Bes. bemühte sie sich um die Erforschung der Bevölkerungsverhältnisse, zunächst ohne explizite Heranziehung der → Wahrscheinlichkeitsrechnung. Ihre heutige Bedeutung verdankt die S. der sich seit 1880 anbahnenden Verbindung mit der Wahrscheinlichkeitsrechnung, deren erste Grundlagen von B. → Pascal und P. de → Fermat um die Mitte des 17.Jahrh. erarbeitet wurden, die im 18.Jahrh. eine Blüte erlebte (Jakob → Bernoulli, A. de → Moivre, L. → Euler, Th. → Bayes) und ihre spätere Form gerade durch die Verbindung mit der S. gewann. Um die Popularisierung der modernen S. machte sich in der Mitte des 18.Jahrh. bes. L.A.J. → Quetelet verdient. Im 19.Jahrh. entstanden drei Richtungen, eine russische, hauptsächlich von P.L.Tschebyschew (1821-94), eine englische, von F. → Galton und eine deutsche, von W. → Lexis begründet. Als zukunftsstärkste Richtung erwies sich die englische. Ihr wichtigster Vertreter war R.A. → Fisher. Die von ihm im wesentlichen zwischen 1920 und 1935 entwickelte Methodik ist die bis heute vorherrschende. Der testtheor. Erweiterung der Fisherschen Methode widmeten sich J.→ Neyman und E.S.→Pearson. In den vierziger Jahren trat zu den klassischen statist. Zielsetzungen (Beschreibung und Analyse) eine weitere, von A. → Wald begründete Zielsetzung hinzu, die Entscheidung unter Risiko und Ungewißheit.

So weit das Zitat aus der Brockhaus Enzyklopädie. Aber auch im so genannten täglichen Leben begegnet uns der Begriff Statistik immer wieder, wie z.B. bei

„... nun kommt die Tabelle der Fußball-Bundesliga für die Statistiker ...“

„... die Bundesanstalt für Arbeit meldet die Arbeitslosenstatistik ...“

„... im statistischen Mittel rauchen Deutsche 18.3 Zigaretten pro Tag ...“

Auf diese fälschliche (volkstümliche) Interpretation der Statistik wies schon *Schott (1914)* hin:

... dass heute unter Statistik allgemein das Ergebnis irgendeiner Zählung verstanden wird. Allein mit solcher Feststellung werden wir unsre Neugier schwerlich als befriedigt erachten wollen; es kann doch nicht sein, dass im Zusammenzählen irgendwelcher Dinge, beliebigen Sinns oder Unsinns, sich schon die Aufgabe der Statistik erschöpft. Auch wenn wir dem Unsinn den Zutritt versperren und als Statistik nur das Ergebnis einer Zählung gelten lassen, die unsre Erkenntnis auf irgendeinem Gebiet bereichert, ist der Kreis noch viel zu weit gezogen, denn wohl gibt es nach unsrer heutigen Auffassung ohne Zäh-

len keine Statistik, aber ebensowenig entsteht diese durch bloßes Zusammenzählen. Der Knabe, der die Äpfel im Frühstückskorb und das Mädchen, das beim Stricken die Maschen zählt, sind dadurch noch nicht zu Statistikern geworden.

Gerade aber in dem Bereich der Sozialarbeit und -pädagogik wie auch in der Pflege ist es wichtig über das reine Zählen hinaus, die jeweilige Situation oder das Umfeld genau zu beschreiben und zu analysieren. Dies wurde auch schon (in leicht abgewandelter Form) von *Schott* (*1914*) ausgedrückt:

Mensch und Menschenwerk sind das Objekt der Statistik im engeren Sinne, oder wie man auch wohl zur Vermeidung von Mißverständnissen zu sagen liebt, der sozialen Statistik. In diesem bescheidener umgrenzten Wirkungskreis der Statistik werden sich unsre weiteren Ausführungen bewegen, denn hier liegen dauernde, im Wesen des Objekts wurzelnde Aufgaben der Statistik vor, die keine Einsicht höherer Art, wie so häufig in den Naturwissenschaften, jemals entbehrlich zu machen vermag. Im Gegenteil: je reicher und vielgestaltiger mit zunehmender Kultur, mit steigender Vergesellschaftung die Beziehungen von Mensch zu Mensch werden, je verwickelter und unübersichtlicher die menschlichen Zwecksetzungen und Willens-äußerungen sich gestalten, desto mehr schwindet die Hoffnung auf einfache Formulierung der Erscheinungen des Gesellschaftslebens. Dieselbe Entwicklung macht aber auch die Bemühungen um einen Weltkatalog in dem oben erwähnten Sinne immer aussichtloser. Die unendlichen Verschiedenheiten der Vorgänge des Gesellschaftslebens einzeln festzuhalten, ist ebenso unmöglich, wie der Versuch, sie auf ausnahmslos gültige Formeln zu bringen, ohne Ertrag bleibt. So muss die Statistik den Retter in der Not abgeben und unter Verzicht auf Allgemeingültigkeit oder Vollständigkeit nach einzelnen wichtigen Merkmalen ihre Gruppen bilden, deren Besetzung ermitteln und so ihr skizzenhaftes Bild entwerfen.

Ähnlich befand auch *Quetelet* (*1914*), in dem er in seinem Vorwort sich folgendermaßen über die Statistik äußerte: ... der vierte [Abschnitt] enthält meine Gedanken über die Lehre vom *mittleren Menschen* und über die Organisation des sozialen Systemes.

Im Rahmen dieses Buches möchten wir darauf eingehen, wie man
- Zahlen festhält,
- deren Typisches, Beachtenswertes, Auffälliges ermittelt,
- ihre verborgenen Regeln und Gesetzmäßigkeiten sowie die jeweiligen Zusammenhänge ermittelt.

Im Wesentlichen können wir uns der Aussage von *Riedwyl* (*1992*) anschließen, der meint: Unter Statistik verstehen wir das Sammeln und Interpretieren von Zahlen, um Hypothesen zu finden und zu überprüfen.

1.3 Teilgebiete der Statistik

Das Gebiet der Statistik lässt sich in (zumindest) drei Teilgebiete untergliedern. Nämlich in die
- deskriptive Statistik,
- induktive Statistik,
- (explorative) Datenanalyse.

Die deskriptive (beschreibende) Statistik befasst sich mit der (komprimierten) Aufarbeitung von Zahlen- oder Datenmaterial. Mit Hilfe der Methoden der deskriptiven Statistik soll das Datenmaterial beschrieben werden. Dazu werden auch vielfach grafische Darstellungen verwendet. Man kann sie als Vorstufe zur induktiven Statistik ansehen.

Die induktive (schließende) Statistik dient dazu, sinnvolle Aussagen zu machen über Auffälligkeiten in gegebenem Datenmaterial, das (zum Teil) nur indirekt oder aber auch nur unvollständig beobachtet wurde. Mit Hilfe der deskriptiven Statistik sollen substanzwissenschaftliche Hypothesen aufgestellt werden, die dann mit Hilfe der induktiven Statistik verifiziert bzw. falsifiziert werden sollen. Zur Durchführung der Methoden und Verfahren der induktiven Statistik wird zumeist ein statistisch-mathematisches Modell benötigt.

Die (explorative) Datenanalyse kann als eine Art Zwitter zwischen deskriptiver und induktiver Statistik angesehen werden. Zum einen versucht sie (wie die deskriptive Statistik), die Daten geeignet zu beschreiben, zum anderen sollen Entscheidungen getroffen werden, ohne auf die Existenz eines statistisch-mathematischen Modells zurückgreifen zu müssen. Sie versucht ihre Schlussfolgerungen daten- und nicht modellorientiert zu ziehen.

Im Rahmen dieses Buches werden wir uns hauptsächlich auf die deskriptive und induktive Statistik beschränken. Falls Methoden oder Verfahren der (explorativen) Datenanalyse vorgestellt werden, wird dies explizit erwähnt.

1.4 Statistik und EDV

Statistik ist heutzutage ohne den Einsatz von EDV nicht mehr vorstellbar. Dies bezieht sich aber nicht nur auf die statistischen Analysen, sondern auch auf die Datenerhebung. In vielen Einrichtungen findet heutzutage eine (gesetzlich vorgeschriebene) Dokumentation statt – unabhängig davon, ob dies EDV-gestützt oder traditionell noch auf Papierbasis geschieht. Dabei sollen die Dokumentation und damit auch die erhobenen Daten, wie z.B. in *Garms-Homolová & Niehörster (1997)* beschrieben, nicht nur dem Leistungsnachweis und zur rechtlichen Absicherung genutzt werden, sondern auch zur Förderung des Informationsflusses, als Kommunikationsmedium, als kontrollierbarer Nachweis und als Organisationsmittel. Doch ohne eine adäquate statistische Analyse der Daten aus der Dokumentation für die jeweilige Problemstellung sind diese Zusatzfunktionen der Dokumentation nicht nutzbar. Um eine Einrichtung evaluieren zu können oder die Dienstleistung einer Einrichtung an die Bedürfnisse ihrer Kundschaft, ihres Klientels oder ihrer Patienten besser ausrichten zu können, bedarf es oftmals einer soliden Datengrundlage. Datenerhebungen machen jedoch keinen Sinn, wenn diese Daten nicht anschließend genutzt, d.h. adäquat ausgewertet und präsentiert werden und dies kann nur unter dem Einsatz von EDV geschehen. Als Beispiele für diese Zusatzfunktionen seien nur kurz erwähnt:

- Erstellung eines Datenblattes für die Zusammenarbeit mit dem MDK
- Dokumentation der Arbeits- und Leistungsnachweise pro Mitarbeiter(in) für regelmäßig stattfindende Personalgespräche
- Reports als unterstützendes Medium bei der Dienstübergabe

- Kennzahlen als Hilfsmittel im Controlling
- Dokumentation und Analyse der Umsetzung von integrierter Versorgung in Netzwerken, von Entlassungsmanagement und Überleitung

Aus diesem Grunde erscheint es als unumgänglich, dass bei der Qualifikation für leitendes Personal (vgl. auch *Ostermann & Wolf-Ostermann (2004)*) eine Grundausbildung in EDV und Empirie bzw. Statistik stattfindet. Es soll jedoch dabei betont werden, dass es dabei nicht darum geht, diese Methodenausbildung überzubetonen. Vielmehr geht es darum, eine Sensibilisierung für diese Problematik zu erreichen, damit die angehenden Leitungskräfte in der Lage sind, entsprechende Statistiken interpretieren zu können. Des Weiteren sollen sie aber auch ihre persönlichen Grenzen in diesem Bereich erfahren, um entscheiden zu können, dass bei höherwertigen Fragestellungen eine externe Beratung herangezogen werden sollte.

2 Empirische Forschung und Statistik

Sowohl im Bereich der Sozialen Arbeit als auch der Pflege sind empirische Untersuchungen unverzichtbare Bestandteile der Forschung. Aufbauend auf erfahrungswissenschaftliche Untersuchungen sollen hierbei Problemlösungsprozesse für ganz bestimmte Fragestellungen entwickelt werden. Methoden, die hierfür Verwendung finden, lassen sich grob in die Bereiche der quantitativen und der qualitativen Verfahren einteilen. Diese beiden Bereiche sollten dabei nicht als konkurrierende Verfahren sondern stattdessen als sinnvolle Ergänzungen zueinander gesehen werden. Im Folgenden soll der Ablauf eines empirischen Forschungsprozesses im Bereich der quantitativen Forschungsmethoden kurz skizziert werden (zur Darstellung qualitativer Forschungsabläufe vgl. etwa *Bortz & Döring (2002)*).

2.1 Ablauf eines empirischen Forschungsprozesses

Ziel jeder erfahrungswissenschaftlichen Untersuchung ist es, für eine bestimmte Fragestellung fundierte Lösungen oder Lösungsansätze zu ent-wickeln. Der Ablauf eines solchen Prozesses lässt sich grob in die folgenden Teilschritte gliedern:

- Definition des Forschungsthemas / der Fragestellung
- Festlegung der Zielgruppe (Grundgesamtheit)
- Festlegung der Untersuchungsgesamtheit
- Entwicklung des Forschungsinstrumentariums (Versuchs-, Erhebungsdesign, Fragenbogendesign, …)
- Anwendung des Forschungsinstrumentariums (Erhebung/Befragung)
- Informationsauswertung (Datenmanagement, statistische Analyse)
- Dokumentation / Präsentation der Ergebnisse

Für die Festlegung des Forschungsthemas ist zunächst eine genaue und klare Abgrenzung der Fragestellung wichtig. Ohne eine solche Präzi-sion ist nachfolgend keine sinnvolle Bearbeitung der Fragestellung möglich. Dies beinhaltet, dass in diesem ersten Stadium einer Untersuchung entweder bereits ein fundiertes Vorwissen vorhanden sein muss und/oder ausgedehnte (Literatur-)Recherchen notwendig sind. Nicht zuletzt impliziert die präzise Formulierung einer Fragestellung auch eine notwendige semantische Präzision. Mangelnde Sorgfalt bei der genauen Festlegung der interessierenden Fragestellung resultiert ansonsten in der Regel in zwei Situationen, die beide gleichermaßen unbefriedigend sind:

- ungenügende Daten zur Beantwortung der Fragestellung („Datenmangel“), d.h. die Fragestellung kann mit den vorliegenden Daten nicht bearbeitet werden
- viele erhobene Daten, die jedoch nichts zur Beantwortung der Fragestellung beitragen können („Datenfriedhof“)

Beide Situationen bedeuten einen erheblichen Aufwand und Kosten, produzieren jedoch nicht die angestrebten Ergebnisse und sollten von daher bereits im Anfangsstadium einer empirischen Untersuchung möglichst ausgeschlossen werden.

Ebenso wichtig wie die Festlegung des Forschungsthemas ist die Festlegung der Zielgruppe (Grundgesamtheit) sowie der Untersuchungsgruppe. Auch hier ist eine klare Abgrenzung wichtig. Mit der Festlegung der Zielgruppe wird definiert, worüber oder über wen nachfolgend Aussagen getroffen werden sollen. Die Untersuchungsgruppe beschreibt i.d.R. eine Teilmenge der Zielgruppe, an der dann tatsächlich die Erhebung durchgeführt wird. Es kann jedoch auch die gesamte Zielgruppe identisch mit der Untersuchungsgruppe sein. Die Auswahl der Untersuchungsgruppe erfolgt anhand theoretischer Kriterien (→ Stichprobenverfahren) sowie praxisbezogener Überlegungen (z.B. Kosten, Erreichbarkeit, …).

Bei der Entwicklung des Forschungsinstrumentariums muss zunächst geklärt werden, welche „Methode“ Verwendung finden soll. Dies ist im Regelfall durch die Fragestellung bereits vorgegeben (*„Die Fragestellung bestimmt die Methode.“*). Einige der bekanntesten Methoden sind:

- die Sekundäranalyse, d.h. die Nutzung bereits vorhandener Daten. Dies ist oftmals eine kostengünstige Analyseform, jedoch „passen“ die Daten nicht immer genau zur Fragestellung, da sie unter einer anderen Zielsetzung erhoben worden sein können.
- das Experiment. Hierbei handelt es sich um eine stark strukturierte Situation („Versuchsplan“), die überwiegend in naturwissenschaftlichen, technischen, psychologischen oder medizinischen Bereichen zur Anwendung kommt.
- die Beobachtung. Hierbei werden anhand eines „Beobachtungsplans“ Daten erhoben. Beobachtungsstudien sind typisch für sozial- und pflegewissenschaftliche Anwendungen.
- die (standardisierte) Befragung. Diese kann sowohl schriftlich, telefonisch, oder persönlich durchgeführt werden und gehört ebenfalls zu den typischen Forschungsmethoden im sozial- und pflegewissenschaftlichen Bereich.

Im Folgenden wollen wir auf die verschiedene Formen standardisierter Befragungen eingehen und deren Vor- und Nachteile kurz vorstellen.

Als Grundformen der standardisierten Befragung lassen sich im Wesentlichen vier Grundformen unterscheiden: mit oder ohne eine befragende Person (Interviewer) sowie traditionellere und an modernere Techniken geknüpfte Verfahren (vgl. Tabelle 2.1):

Form	mit Interviewer	ohne Interviewer
„traditionell“	persönlich/mündlich (PAPI/CAPI[2])	postalisch
„moderner“	telefonisch (CATI[3])	per Internet

Tabelle 2.1 Formen der standardisierten Befragung

Die persönlich/mündliche Form der Befragung unter zu Hilfenahme eines Interviewers ist das klassische Instrument zur Bevölkerungsbefragung und findet in vielen Bereichen seine Einsatzmöglichkeiten, z.B. der Volkszählung (Mikrozensus), allgemeiner Meinungsforschung, Wahlprognosen, Mitarbeiter-/Patientenbefragungen, etc. Auch im Bereich der Sozialen Arbeit ist an vielen Stellen der Einsatz von Interviewern zu empfehlen. Die Auswahl der zu Befragenden erfolgt nach einem ausgearbeiteten „Stichprobenplan“. Durch die direkte Form der Befragung ist es möglich, nachzufragen, zu motivieren, Fragen und/oder Antwortvorgaben zu erläutern und auch Hilfsmittel (Bilder, Gegenstände, etc.) einzusetzen. Die Zeit, die zur Beantwortung eines Fragebogens zur Verfügung steht, ist relativ kurz und es ist keine Kontrolle / Rücksprachemöglichkeit des Interviewers gegeben. In den letzten Jahren haben moderne EDV-gestützte Fragebögen mehr und mehr an Bedeutung gewonnen. Zum einen wird dabei versucht, einen Papier-Fragebogen so zu gestalten, dass er direkt mit einem Belegleser in einen Computer eingelesen werden kann, um so die Fehlerrate gegenüber einer Übertragung per Hand zu verringern (siehe auch *Kreienbrock & Ostermann (1993)*). Zum anderen werden Fragebögen direkt auf einem Computer-Bildschirm (Laptop) der zu befragenden Person gezeigt, um jegliche Übertragungsfehler zu vermeiden. Letztere Methode ist unter dem Stichwort CAPI (Computer Assisted Personal Interview) bekannt. Im Bereich der Sozialen Arbeit kann jedoch der Einsatz eines CAPI zu unvermuteten Schwierigkeiten führen, da z.B. bei psychisch Kranken eine hohe Ablehnung gegenüber diesem Instrumentarium zu verzeichnen ist.

Ein weiteres klassisches Instrument zur Bevölkerungsbefragung ist die telefonische Form der Befragung (ebenfalls unter zu Hilfenahme eines Interviewers). Telefonbefragungen haben oftmals den Vorteil, dass sie preisgünstiger sind als der Einsatz von Interviewern und einen hohen Zeitgewinn bieten. Sie werden eingesetzt für so genannte „Blitzumfragen“, zur Meinungsforschung, etc. Die zuvor gemachten Anmerkungen zum persönlichen Interview gelten im Wesentlichen auch für diese Befragungsform. Nicht möglich ist natürlich der Einsatz von Hilfsmitteln (mit Ausnahme von akustischen Hilfsmitteln). Die Beantwortungszeit ist sehr kurz und erfolgt unter einem gewissen „Zeitdruck“. Es besteht jedoch eine Kontroll-/ Rücksprachemöglichkeit des Interviewers.

Beide Befragungstypen setzen eine situationsgerechte Konstruktion des Befragungsbogens (einfache Fragen, nicht zu viele Antwortmöglichkeiten, …) voraus. Sehr wichtig ist auch eine sorgfältige Schulung der Interviewer, da hiervon der Erfolg der Umfrage ganz wesentlich

[2] PAPI: Papfer Assisted Personal Interview / CAPI Computer Assisted Personal Interview

[3] CATI Computer Assisted Telephone Interview

abhängt. Dies bedeutet zum einen, dass der Interviewer zuverlässig ist, sich z.B. genau an den vorgegebenen Auswahlplan hält, so dass keine Verfälschungen (etwa aus Bequemlichkeit) auftreten. Zum anderen müssen aber auch Auftreten, Stimme und nicht zuletzt auch Kleidung der Befragungssituation angepasst sein. Neben vielen positiven Effekten einer Interviewer-gestützten Befragung ist jedoch auch zu berücksichtigen, dass diese in aller Regel mit großen Kosten (da sehr personalintensiv) verbunden ist.

Vielfach werden deshalb Umfragen ohne den Einsatz von Interviewern durchgeführt. Auch bei „anonymen" Befragungen erfolgt die Auswahl der Befragten prinzipiell nach einem genauen „Stichprobenplan". Da die Fragen und Antwortmöglichkeiten vom Befragten selbst gelesen werden, sind sehr klare Formulierungen und evtl. schriftliche Erläuterungen im Befragungsbogen notwendig. Der Einsatz von Hilfsmitteln ist bei dieser Form der Befragung nur sehr bedingt möglich. Prinzipiell ist die Beantwortungszeit beliebig lang, eine Kontrolle der Befragungssituation (wer antwortet und in welcher Antwortreihenfolge) ist nicht mehr möglich. Bei der postalischen Form einer solchen Befragung ist es zudem sehr wichtig, einen Anreiz zur Rücksendung des ausgefüllten Fragebogens – etwa durch eine bereits adressierten und frankierten Rückumschlag und/oder die Koppelung mit einem Preisausschreiben o.ä. – zu liefern, um einen zufrieden stellenden Rücklauf der Fragebögen sicherzustellen. In aller Regel erfolgt eine starke Selbstauswahl der Befragten - trotz eines vorher aufgestellten Stichprobenplanens – da oftmals nur Personen mit einem starken Interesse an dem Thema (positiv wie negativ) zu einer Antwort bereit sind.

Bei einer Internet-basierten Befragung ist die prinzipielle Auswahl der Befragten nach einem Stichprobenplan nicht möglich, sondern die Selbstauswahl der Befragten die Regel. Lediglich durch das Setzen von „Links" auf bestimmten Internetseiten kann versucht werden, die Auswahl der Befragten ansatzweise zu steuern. Hierbei ist auch darauf hinzuweisen, dass bei Internet-basierten Befragungen evtl. bestimmte Bevölkerungsgruppen nicht erreichbar oder unterrepräsentiert sind (Ältere, Frauen etc.) Befragungen im Internet können jedoch gerade im Kinder- und Jugendbereich eine interessante Alternative sein. Positiv bei Internetbefragungen ist zu sehen, dass durch den Einsatz von (programmierten) Filtertechniken Hinweise auf fehlende oder falsche Eingaben gegeben werden können. In allen Fällen einer schriftlichen Befragung ist der Kostengesichtspunkt nicht zu vernachlässigen. Bei Internet-basierten Befragungen kommt hinzu, dass der Aufwand für die technischen Voraussetzungen, den Programmieraufwand, Fragen der Datensicherheit und Anonymität nicht unterschätzt werden sollte.

Insgesamt gilt für alle Typen einer standardisierten Befragung, dass Gesichtspunkte der praktischen Durchführung von vornherein in einem angemessenen Ausmaß bei der Planung der Befragung berücksichtigt werden müssen. Dazu gehören die (teilweise schon angesprochenen) Aspekte bei der Formulierung der Fragen, von Ausfüllanweisungen und Anschreiben und/oder Erinnerungsschreiben oder aber die Organisation von Versand und Rücksendung, mögliche Interviewerschulungen und natürlich die Berücksichtigung des zur Verfügung stehenden Budgets. Wichtig ist bei allen Planungen, dass eine sorgfältige Dokumentation des Ablaufes und insbesondere von Abweichungen von der ursprünglichen Erhebungsplanung erfolgt. Daneben muss natürlich auch sichergestellt sein, dass notwendigen Formalitätsaspekten Rechnung getragen wird:

- Nennung von Institution / Auftraggeber der Befragung
- je nach Umfrageart: persönliche Vorstellung des Interviewers, Erklärung von Auswahlrichtlinien, Angaben zur Befragungsdauer,…
- Erläuterung des Themas
- Zusicherung (und Einhaltung!) von Anonymität / Datensicherheit[4]
- Hinweis auf Freiwilligkeit bzw. Einholen des Einverständnisses von Erziehungsberechtigten bei Minderjährigen (Richtlinien einhalten!)

2.2 Fragebogendesign

Bei der Konstruktion des Fragebogens selbst ist darauf zu achten, dass sowohl die Einzelfragen als auch der gesamte Bogen gewissen Gütekriterien entsprechen. So muss der Fragebogen als Ganzes insbesondere die üblichen Gütekriterien

- Objektivität: Durchführung, Auswertung und Interpretation sind unabhängig von der anwendenden Person
- Reliabilität: wie groß sind Fehlereinflüsse? (Messgenauigkeit!)
- Validität: wird tatsächlich gemessen, was gemessen werden soll?

erfüllen. (Für eine genauere Definition der Gütekriterien und weitere Ausführungen hierzu vgl. etwa *Bortz & Döring (2002, S.193ff)*). Dies bedeutet für die einzelnen Fragen und Fragenkomplexe eines Bogens, dass zunächst genau über die Zielsetzung der jeweiligen Fragen nachgedacht werden muss, um dann zu einer sorgfältigen zielführenden Formulierung der Fragen zu gelangen. Hierbei sollten u.a. die nachfolgend aufgelisteten Kriterien berücksichtigt werden:

- Relevanz bzgl. Fragestellung
- Abdeckung des gesamten affektiven Bereichs
- Vermeidung „allgemeiner“ Sachverhalte / Tatsachenbeschreibungen
- keine Suggestivfragen
- Formulierung eines Gedankens pro Frage
- kurze, klare, direkt verständliche Formulierungen
- Verwendung positiver und negativer Formulierungen
- eindeutige Interpretation

Auch die Anzahl vorgegebener Antwortkategorien beeinflusst die später damit erzielten Ergebnisse. Bei abgestuften (ordinalen) Antwortkategorien (z.B. *gefällt mir nicht, teils/teils, gefällt mir*) führt eine gerade Anzahl vorgegebener Kategorien dazu, dass die Befragten sich jeweils positionieren müssen, wohingegen eine ungerade Anzahl von Kategorien in aller Regel eine „neutrale“ (mittlere) Position erlaubt. Bei nicht abgestuften (nominalen) Antwortkategorien muss prinzipiell über das Problem der Reihenfolge (alphabetisch? nach Wichtigkeit?) nachgedacht werden. Zudem sollte darauf geachtet werden, dass nicht zu viele Kategorien vorgegeben werden, da dies dazu führen kann, dass nur noch die ersten (und evtl. die

[4] siehe Bundesdatenschutzgesetz

letzten) Kategorien wirklich Beachtung finden. Für weitere Ausführungen sei etwa auf *Schwarz et al. (1989)* verwiesen. Bei der Verwendung von Alternativfragen (nur zwei Antwortkategorien sind möglich, etwa ja/nein) sollten jeweils beide Kategorien auf dem Bogen sichtbar sein, damit nachfolgend auch fehlende Werte noch unterschieden werden können.

Neben einer Einteilung von Fragen nach Form der Antwortvorgaben in

- offene Fragen: haben ein breites Antwortspektrum (Eindimensionalität?), jedoch müssen die Antworten i.d.R. nachträglich strukturiert werden und
- geschlossene Fragen: vorgegebene Antwortkategorien (Alternativfragen / Auswahlfragen)

können Fragen auch nach ihrem Zweck klassifiziert werden (vgl. auch *Becker (2004)*):

- Einleitungs- /Übergangsfragen stellen Bezugsrahmen her
- Filterfragen dienen dem Ausschluss nicht relevanter Personengruppen
- Hauptfragen erfragen die wesentlichen Sachverhalte (demografische Fragen, Fakt-, Meinungs-, Beurteilungs- und Verhaltensfragen)
- Folgefragen dienen der genaueren Analyse von Antworten
- Kontrollfragen ermöglichen eine Antwortkontrolle bzw. dienen der Erfassung der Antwortqualität
- indirekte Fragen dienen zur Erfassung „sensibler" Themen
- Abschlussfragen geben dem Befragten die Möglichkeit zu Anmerkungen und Kommentaren

Aber auch die Reihenfolge der Fragen in dem gesamten Bogen sowie das gewählte Layout sind entscheidend für den Erfolg der Umfrage. So ist es etwa entscheidend für welche Klientel der Bogen gedacht ist:

- muss eine besondere Schriftgröße z.B. für Hochbetagte oder sehbehinderte Personen gewählt werden?
- sind anstelle von schriftlich formulierten Antwortkategorien Symbole – z.B. „ ☺ " – vorzuziehen, etwa bei Grundschulkindern, etc.?

Generell gilt, dass zu kleine oder ausgefallene Schrifttypen zu vermeiden sind, Farben eher dezent Verwendung finden sollten und der Bogen insgesamt ein ausgewogenes und dem Zweck angemessenes Layout erhält.

Bevor der entwickelte Fragebogen in einer Untersuchung tatsächlich zum Einsatz kommt, sollte er einem Pretest unterzogen werden. Dieser dient dazu evtl. Unklarheiten (etwa in den Formulierungen) und/oder Probleme in der Anwendung noch vor der großflächigen Anwendung „abzufangen".

Wichtig ist bei allen genannten Punkten, sich zu vergegenwärtigen, dass der Erfolg und die Qualität der wissenschaftlichen Erhebung in hohem Maße von den hier beschriebenen „Vorüberlegungen" und „Vorarbeiten" abhängt und mangelnde Sorgfalt nachfolgend nicht mehr zu korrigieren ist.

Bewusst machen muss man sich aber auch, dass es trotz aller Sorgfalt keine Garantie dafür gibt, dass die gemessene (Einstellungs-) Werte mit den „wahren" (Einstellungs-) Werten tat-

sächlich übereinstimmen. So ist immer zu berücksichtigen, dass jede Umfrage abhängig ist von

- der jeweiligen Situation, der Kooperationsbereitschaft der Befragten
- Verzerrungen aufgrund kognitiver Effekte
- einer Vielzahl möglicher Urteilsfehler (Hawthorne-Effekt, Sponsorship-Bias, Soziale Erwünschtheit, Self-Disclosure, Halo-Effekt, Milde-Härte-Fehler, Tendenz zur Mitte, Primacy-Regency-Effekt, Rater-Rater-Interaktion, ... (vgl. hierzu etwa *Bortz & Döring (2002, S.182ff)*)

Mögliche Strategien zur Vermeidung dieser Problematiken können nur in einer ausgesprochen sorgfältigen Vorbereitung der Umfrage sowie in der ausführlichen Darstellung der Untersuchung und ihrer Ziele und der Zusicherung von Anonymität für die Befragten bestehen.

Die hier angesprochenen Aspekte sollen nur beispielhaft aufzeigen, was für eine gute Fragebogengestaltung wichtig sein kann, ohne dass diese Aufzählung hier vollständig ist. Weitere Ausführungen findet man in den einschlägigen Werken zur empirischen Sozialforschung wie z.B. *Atteslander (2003), Roth (1993)* oder *Holm (1975)*.

2.3 Datenmanagement und Datenanalyse

Nach der Entwicklung eines Fragebogens und der eigentlichen Umfrage ist ein gutes Datenmanagement notwendig, damit die gewonnen Ergebnisse auch adäquat verarbeitet und genutzt werden können. Auch dieser Teilbereich eines empirischen Forschungsprozesses steht noch vor der eigentlichen Auswertung der Ergebnisse und kann sich ebenfalls sehr zeitintensiv gestalten, da auch hier eine große Sorgfalt notwendig ist.

Zunächst sollte schon parallel zur Fragebogenentwicklung über notwendige Kodierungsschemata nachgedacht werden und evtl. auch schon darauf aufbauende Dateneingabemasken entwickelt werden. Hierzu gehören u.a.

- die Vergabe von Fragebogennummern bzw. PIDs (Persönliche Identifikations-Nummern)
- die interne Kodierung von Umfragemerkmalen (wie etwa Ort, Datum, Interviewer, …)
- die (evtl. hierarchische) Kodierung der Antworten

Jeder Fragebogen sollte mit einer (mehrteiligen) Fragebogennummer versehen werden. Diese Fragebogennummer hat oftmals sowohl den Zweck, den einzelnen Bogen zu identifizieren und für spätere Korrekturen des eingegebenen Datensatzes noch einmal zugänglich zu machen, als auch zur internen Kodierung von Umfragemerkmalen zu dienen. Durch eine geeignete Verschlüsselung lassen sich hierbei einige (vornehmlich diskrete) Variablen direkt kodieren, wie etwa:

- Interviewer
- Ort und Datum der Befragung
- Befragungswelle

Die Fragebogennummer muss bei der Dateneingabe ebenfall mit aufgenommen werden. Versieht man den Fragebogen noch vor der Befragung mit einer Nummer, kann dies jedoch evtl. zu Schwierigkeiten führen. Einige der Befragten unterliegen dann oftmals dem Irrtum, dass die ihnen zugesagte Anonymität nicht mehr gewährleistet ist. Man sollte in diesen Fällen entsprechende „Rezepte“ parat haben, und ggf. die Fragebogennummer nachträglich handschriftlich auf den Fragebogen schreiben.

Kodierungsschemata für alle Fragen orientieren sich grundsätzlich daran, dass für jede Antwortkategorie einer Frage auf dem Fragebogen i.d.R. eine Ziffer vergeben wird. Diese Ziffern sollten in sinnvoller Weise die vorgegebenen Kategorien abbilden. So sollte etwa bei gestuften (ordinalen) Kategorien auch die gewählten Ziffern diese Ordnung wiedergeben. Fehlende Werte sollten nicht als leere Datenfelder eingetragen sondern ebenfall explizit kodiert werden. Es scheint nahe zu liegen, bei „Nichtantworten“ den Kode „0“ zu vergeben. Kann jedoch der Wert 0“ auch als tatsächlicher Zahlenwert in einer Umfrage erzielt werden, ist es oftmals besser eine andere, deutlich von tatsächlichen Zahlenwerten abweichende Kodierung zu wählen, etwa den Wert „-999“.

Es kann in manchen Situationen auch sinnvoll sein, die Behandlung der fehlenden Werte umfangreicher vorzunehmen. Manchmal ist man in der glücklichen Lage, dass man die fehlenden Werte selbst in unterschiedliche Kategorien einteilen kann, da für ihr Auftreten unterschiedliche Gründe vorliegen. So könnte man z.B. die Kodierungen

-1 die Person wurde (noch) nicht angetroffen
-2 die Person verweigerte die Auskunft
-3 die Daten gingen verloren

vornehmen, um die oben genannten drei verschiedenen Gründe zu klassifizieren. Man wählt diese Vorgehensweise, da man z.B. bei den Kodierungen –1 und –3 durch eine weitere Befragung doch noch zu den für die Umfrage erforderlichen Daten gelangen kann. Bei der Kodierung –2 wird man wohl davon ausgehen, dass auch bei einem zweiten Versuch die Antwort verweigert wird.

Auf Grund langjähriger Praxiserfahrung ist aber noch ein weiteres Phänomen bekannt: Obwohl nur eine Antwort zugelassen ist, wird mehr als eine Antwort gegeben. So ist es keine Seltenheit, dass selbst bei der Frage nach dem Geschlecht sowohl „weiblich“ als auch „männlich“ angekreuzt wird. In diesen Fällen empfiehlt es sich, einen weiteren Kodewert zu vergeben. Somit hat man dann z.B. beim Geschlecht die Kodierungen

0 keine Angabe
1 weiblich
2 männlich
3 mehr als eine Angabe

Damit trägt man der Tatsache Rechnung, dass fehlende Werte andere Sachverhalte darstellen als nicht erlaubte Mehrfachnennungen.

Nach der Festlegung der Kodierungsschemata kann die eigentliche Dateneingabe erfolgen. Hierbei sollten folgende Punkte beachtet werden:

- Der Datensatz sollte entweder in einer Datenbank oder aber als Einzeldatensatz in Matrixstruktur (je Bogen eine Zeile, je Item einen Spalte) angelegt werden

- Die zuvor festgelegten Kodierungen müssen beachtet werden.
- Die Anzahl von Ziffern und Nachkommastellen muss festlegt werden.
- „Labelvereinbarungen“ für Variablen und Kodierungen
- Zu jedem Datensatz sind Sicherheitskopien anzulegen!

Die Dateneingabe kann zur größeren Genauigkeit (und bei vorhandenen Ressourcen) evtl. doppelt mit nachfolgendem Datenabgleich erfolgen. Abgeschlossen wird jede Dateneingabe immer durch Plausibilitätskontrollen des Datensatzes, um nachfolgend auch valide Analysen durchführen zu können.

Erst daran schließt sich dann die eigentliche Analyse des Datensatzes an. Hierbei erfolgt zunächst eine eindimensionale (überwiegend deskriptive) Analyse der erhobenen Daten. Erst im Anschluss daran finden ausgewählte mehrdimensionale Auswertungsverfahren Verwendung. Im Bereich der induktiven Methoden ist dabei vor einer inflationären Verwendung statistischer Testverfahren zu warnen, da hierbei oftmals das Problem des multiplen Testens und der damit verbundnen mangelnden Gültigkeit von erzielten p-Werten nicht ausreichend berücksichtigt wird. Wie wichtig eine ausreichende Methodenkenntnis zur adäquaten Auswertung von Datensätzen und auch die Kenntnis entsprechender Programmsysteme (SPSS, Excel, …) ist, muss an dieser Stelle sicher nicht explizit erwähnt werden.

2.4 Dokumentation und Präsentation

Als letzter Schritt eines empirischen Forschungsprozesses schließt sich dann die Dokumentation und Präsentation der erzielten Ergebnisse an. Um die erzielten Ergebnis einordnen zu können ist die Beschaffung von zusätzlichen Informationen und Literatur sowie der Umgang damit eine notwendige Voraussetzung. Kenntnisse formaler Aspekte und Sicherheit im Umgang mit diesen formalen Aspekten sind ebenfalls grundlegend. Hierzu gehört auch die Kenntnis entsprechender Textverarbeitungs- und/oder Präsentationsprogramme. Um Ergebnisse aus Auswertungsprogrammen in diese zu übertragen, müssen Importfunktionen für Grafiken und Tabellen ebenso beherrscht werden wie automatische Indizierungen, Querverweise und Formatvorlagen in den Textverarbeitungsprogrammen.

Das Erlernen von Präsentationstechniken (weniger ist hier oft mehr) sowie von Vortragstechniken ist wichtig, um die gewonnen Erkenntnisse auch einer größeren (Fach-) Öffentlichkeit zugänglich zu machen.

Für schriftliche Veröffentlichungen sei hier nur kurz eine beispielhafte Gliederung von empirischen Arbeiten aufgenommen

- Einleitung
- (Theoretische) Einführung in das Thema
- Hypothesen und Modellbildung / Fragestellung
- Material und Methode(n)
- Ergebnisse
- Diskussion

- Ausblick
- Literaturverzeichnis und Anhänge (Verzeichnisse von Abbildungen, Tabellen, Abkürzungen, Fragebogen, …)

Für weitergehende Hinweise hierzu sei auf entsprechende Literatur verwiesen (z.B. *Reinhart (2003), Sesink (1994), Werder, (1993)*).

In den nachfolgenden Kapiteln wollen wir uns aber allgemeinen methodischen Fragestellungen nicht mehr zuwenden und uns auf die methodischen Aspekte statistischer Analysen beschränken.

3 Grundlegende Begriffe

Nachdem in dem vorhergehenden Kapitel kurz auf die Probleme der Datenerhebung eingegangen worden ist, sollen nun grundlegende Begriffe der deskriptiven Statistik (beschreibenden Statistik) definiert bzw. erläutert werden, um eine einheitliche Begriffsbildung zu schaffen.

3.1 Untersuchungseinheiten und Merkmale

Daten, die im Rahmen einer Studie oder einer Untersuchung erhoben werden, misst man an Personen, Tieren oder Objekten. Diese Personen, Tiere oder Objekte werden Untersuchungseinheiten genannt. Die Größen, die an den Untersuchungseinheiten gemessen werden, heißen Merkmale. Die (theoretischen) Werte, die die Merkmale annehmen können, nennt man Merkmalsausprägungen. Oft wird anstelle des Begriffs der Merkmalsausprägung auch der Begriff Realisation verwandt.

Beispiel 3.1 *Werden in einer Schulklasse die Körpergrößen und -gewichte der Kinder gemessen, so sind die Schulkinder die Untersuchungseinheiten, Körpergröße bzw. -gewicht die beiden Merkmale und 125cm, 128cm, 118cm bzw. 28kg, 29 kg, 24kg sind mögliche Merkmalsausprägungen.*

Beispiel 3.2 *Bei einer Umfrage über die Situation von Pflegeheimen könnten die Untersuchungseinheiten die Heime selber sein. Mögliche Merkmale sind Träger, Anzahl der Einzel- bzw. Doppelzimmer, Anzahl Pflegekräfte.*

Beispiel 3.3 *Im Rahmen einer Umfrage an einer Fachhochschule werden die Studierenden nach Alter, Semesterzahl und Studienziel befragt. Hier sind die Untersuchungseinheiten die Studierenden. Die Merkmale sind Alter, Semesterzahl und Studienziel. Mögliche Ausprägungen sind 22, 28, 35 Jahre; 1., 3., 7. Semester sowie Sozialarbeiter, Sozial- oder Heilpädagoge.*

Beispiel 3.4 *In einer Studie über Familienunterstützende Dienste sind die Untersuchungseinheiten die Familien, die diesen Dienst in Anspruch nehmen. Interessierende Merkmale sind z.B. wöchentliche Stundenzahl der Inanspruchnahme, welches Familienmitglied betreut wird, die Höhe der Selbstkostenbeteiligung.*

3.2 Das Skalenniveau von Merkmalen

Merkmale lassen sich auf vielfältige Art und Weise in Klassen oder Niveaus einteilen. Die wichtigsten Einteilungen werden im Folgenden vorgestellt.

Die Skala mit dem niedrigsten Niveau ist die Nominalskala. Die Ausprägungen eines Merkmals, das einer Nominalskala genügt, unterliegen keiner Ordnung. Man kann für die Ausprägungen nur entscheiden, ob sie gleich oder ungleich sind.

Beispiel 3.5 *Im Folgenden sind einige nominalskalierte Merkmale und ihre möglichen Merkmalsausprägungen aufgelistet:*

Beruf	*Sozialarbeiter, Krankenschwester, Altenpfleger, ...*
Geschlecht	*weiblich, männlich*
Farbe	*rot, blau, gelb, violett, ...*
Staat	*Belgien, Deutschland, Dänemark, ...*

Eine Skala mit höherem Niveau ist die Ordinalskala (Rangskala). Für Ausprägungen eines ordinalskalierten Merkmals kann man entscheiden, ob sie gleich oder ungleich sind, und die Ausprägungen lassen sich in einer Reihenfolge anordnen. Obwohl auf der Ordinalskala eine Ordnung vorliegt, sind die Abstände auf dieser Skala nicht interpretierbar. So ist das Merkmal Schulnote ordinalskaliert, jedoch ist ein Schüler, der die Note „gut“ erhalten hat, nicht doppelt so gut wie ein Schüler mit der Note „ausreichend“; er ist nur besser. (Würde man die Schulnoten nicht mit den Zahlen 1, 2, 3, ... in Verbindung bringen, sondern mit den Buchstaben A, B, C, ..., so wäre die Schulnote leichter als ordinalskaliert aufzufassen.) Auf einer Ordinalskala ist es nicht erforderlich, dass die Merkmalsausprägungen konsekutiv benannt werden, also z.B. 1, 2, 3, ... oder A, B, C, ... Es muss jedoch die Reihenfolge (Ordnung) eindeutig erkennbar sein, d.h., man muss entscheiden können, ob zwei Ausprägungen gleich oder ungleich sind, und im Falle der Ungleichheit, welche der beiden Ausprägungen die „größere“ bzw. die „kleinere“ ist. Das Ordnungsprinzip „kleiner“, „gleich“ oder „größer“ ist dabei nicht strikt numerisch zu verstehen, da z.B. auch verbale Beurteilungen – etwa Antwortvorgaben in einem Fragebogen – in eine Reihenfolge gebracht werden können.

Beispiel 3.6 *Im Folgenden sind einige ordinalskalierte Merkmale und einige ihrer möglichen Merkmalsausprägungen aufgelistet:*

Schulnote	*1, 2, 3, 4, 5 und 6*
Note beim Eiskunstlaufen	*6.0, 5.9, 5.8, ...*
Handelsklassen	*A, B, C, ...*
Beurteilung (Fragebogen)	*gefällt mir, weiß nicht, gefällt mir nicht*
Pflegestufe	*I, II und III*

Die Skala mit dem höchsten Niveau ist die metrische Skala (Kardinalskala). Auf ihr ist nicht nur eine Ordnung gegeben, sondern die Abstände sind interpretierbar. Für Ausprägungen eines metrischskalierten Merkmals kann man daher entscheiden, ob die Ausprägungen gleich oder ungleich sind, man kann sie in eine Reihenfolge bringen und diese Reihenfolge auch bezüglich ihrer Abstände und evtl. auch bzgl. ihrer Verhältnisse zueinander interpretieren.

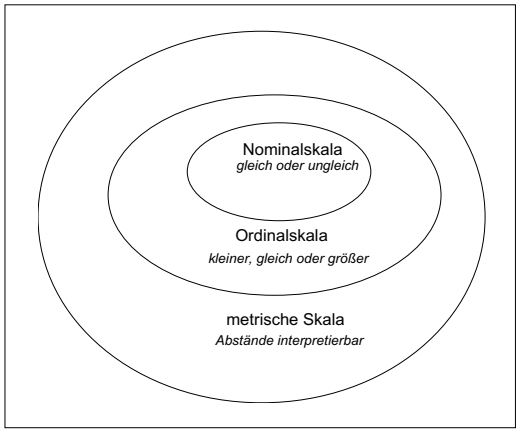

Abbildung 3.1 Schematische Darstellung der verschiedenen Skalenniveaus und zugehörige Interpretationsmöglichkeiten der Ausprägungen

Die metrische Skala wird zusätzlich unterteilt in die Intervallskala und in die Verhältnisskala. Eine Verhältnisskala besitzt einen natürlichen Nullpunkt (Ursprung), eine Intervallskala einen künstlichen. So ist eine Temperaturskala in Celsiusgraden eine Intervallskala, da der Nullpunkt (Schmelzpunkt Eis → Wasser) willkürlich gewählt wurde. Eine Temperaturmessung in Kelvin entspricht einer Verhältnisskala, da Null Grad Kelvin beim absoluten Nullpunkt gemessen werden. Der Unterschied zwischen den beiden Skalen soll an diesem Beispiel noch einmal erläutert werden:

- Stehen in einem Berliner Pflegeheim zwei Tassen mit Tee, wobei der Tee der ersten Tasse eine Temperatur von 40 Grad Celsius und die der zweiten von 20 Grad Celsius, so darf man dennoch nicht behaupten, dass der Tee in der ersten Tasse doppelt so heiß sei wie der Tee der zweiten Tasse. Denn würden die beiden Tassen in New York stehen, so hätten der Tee der ersten Tasse eine Temperatur von 104 Grad Fahrenheit und der Tee der zweiten Tasse eine Temperatur von 68 Grad Fahrenheit. Das Verhältnis von 104 zu 68 beträgt nun nicht mehr 2:1.
- Da es sich aber in beiden Fällen um eine Intervallskala handelt, darf man nun sagen: Die Temperatur des Tees der ersten Tasse ist doppelt so weit von Schmelzpunkt des reinen Wassers (mit 32 Grad Fahrenheit) entfernt wie die Temperatur des Tees der zweiten Tasse, denn auch bei Messung in Fahrenheit gilt:

$$\frac{104-32}{68-32}=\frac{72}{36}=\frac{2}{1} \quad .$$

Beispiel 3.7 *Im Folgenden sind einige metrischskalierte Merkmale und die Angabe, ob sie einer Intervall- oder einer Verhältnisskala entsprechen, aufgelistet:*

Körpergröße	*Verhältnisskala*
Übergewicht (in kg)	*Intervallskala*
Pflegemehraufwand (in Minuten)	*Intervallskala*
Anzahl der Seiten in einem Buch	*Verhältnisskala*

Vielfach werden Merkmale auch in qualitative und quantitative Merkmale untergliedert. Die qualitativen Merkmale unterscheiden sich durch ihre Art, die quantitativen durch ihre Größe. Nominale Merkmale sind daher immer auch qualitative Merkmale und metrischskalierte Merkmale sind immer quantitative Merkmale. Ordinale Merkmale werden entweder den qualitativen Merkmalen zugeordnet oder aber unter dem Begriff der komparativen (vergleichenden) Merkmale erfasst.

Außerdem wird noch in diskrete und in stetige Merkmale unterschieden. Diese Einteilung ist jedoch nur sinnvoll, wenn den Merkmalen eine metrische Skala zugrunde liegt. Diskrete Merkmale besitzen nur endlich viele mögliche Ausprägungen, stetige Merkmale unendlich viele (bei beliebig feiner Messgenauigkeit). Der Übergang von diskreten zu stetigen Merkmalen ist nicht genau festgelegt, darum spricht man oftmals bei einem diskreten Merkmal mit sehr vielen möglichen Ausprägungen von einem approximativ stetigen Merkmal. Zudem kann die Frage, ob ein metrisches Merkmal diskret oder stetig ist, auch von dem jeweiligen Kontext abhängen, in dem das Merkmal beobachtet wird.

Beispiel 3.8 *Das Merkmal „Kinder ja/nein" ist qualitativer Natur, die Angabe der „Kinderzahl“ (0, 1,2,...) dagegen ein quantitatives und diskretes Merkmal. Die (beliebig genau) gemessene Körpergröße ist ein quantitatives und stetiges Merkmal. Das quantitative Merkmal „Alter“ z.B. würde man bei einer Untersuchung im Kindergarten mit den Ausprägungen 3, 4, 5 und 6 als diskretes Merkmal auffassen, aber bei einer Untersuchung über die bundesrepublikanische Bevölkerung mit einem Wertebereich von 0 bis 120 Jahren als approximativ stetiges oder sogar als stetiges Merkmal.*

3.3 Grundgesamtheit versus Stichprobe

Wir messen Merkmale an Untersuchungseinheiten und interessieren uns für ihre Merkmalsausprägungen. Dabei stellt sich die Frage, woher diese Untersuchungseinheiten stammen. Dazu ist es notwendig, die Begriffe Grundgesamtheit und Stichprobe zu definieren. Unter einer Grundgesamtheit versteht man alle Untersuchungseinheiten, denen das zu untersuchende Merkmal gemeinsam ist. Eine Grundgesamtheit lässt sich nach räumlichen, zeitlichen und/oder sachlichen Gesichtspunkten eindeutig charakterisieren bzw. abgrenzen. Diese Abgrenzung ergibt sich im Einzelfall immer aus der jeweiligen Problemstellung. Den Umfang der Grundgesamtheit bezeichnen wir mit N. Werden alle Untersuchungseinheiten in die Erhebung einbezogen, d.h., wird an allen Untersuchungseinheiten der Grundgesamtheit das interessierende Merkmal gemessen, so spricht man von einer Voll- oder Totalerhebung. Das bekannteste Beispiel einer (Beinah)-Totalerhebung ist die Volkszählung. Volkszählungen gab es im Altertum schon bei den Chinesen, Ägyptern und Juden (2.Sam. 24, 1-10). In Deutschland wurden schon früh Stadt- bzw. Landeszählungen durchgeführt (Nürnberg 1449, Preußen 1725). Nach Vorschlägen der UNO sollen Volkszählungen in heutiger Zeit etwa alle 10 Jahre stattfinden. Volkszählungen dienen dazu, staatlichen Behörden und Einrichtungen Informationen zu liefern, damit diese aufgrund der Datenbasis der Volkszählungen Planungen durchführen können. So ist z.B. eine Planung und Abschätzung des Bedarfes an Kindergartenplätzen ohne derartige Datenerhebungen nicht möglich.

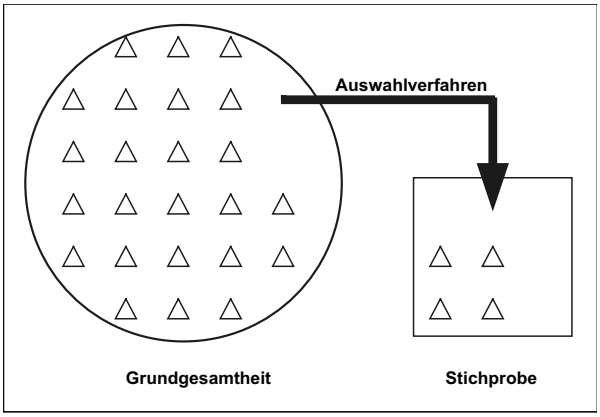

Abbildung 3.2 Vorgehensweise zur Erstellung einer Stichprobe

Totalerhebungen, also die Erfassung der vollen Grundgesamtheit, stellen unter dem Gesichtspunkt einer statistischen Auswertung die wünschenswerteste Form der Datenerhebung dar. In der Regel ist eine Totalerhebung jedoch aus zeitlichen, finanziellen oder anderen Gründen nicht durchführbar. Aus diesem Grund wird versucht, eine Teilmenge aus der Grundgesamtheit zu erfassen, die es ebenfalls ermöglichen soll, sinnvolle Aussagen über die Grundgesamtheit zu treffen.

Eine Teilmenge von Untersuchungseinheiten aus einer Grundgesamtheit bezeichnen wir als Stichprobe. Ihren Umfang bezeichnen wir mit n. Es gilt: $n \leq N$. Bei einer Stichprobenuntersuchung spricht man auch von einer Teilerhebung. Die bekannteste Teilerhebung ist der Mikrozensus. Der Mikrozensus dient der Aktualisierung der Volkszählungsergebnisse. In der Bundesrepublik Deutschland werden vierteljährlich 0.1% und jährlich 1% aller Haushalte im Rahmen des Mikrozensus befragt.

Auf Grund des jeweilig verwandten Auswahlverfahrens werden die Stichproben verschiedenartig bezeichnet. Hierbei kann man z.B. unterscheiden, ob ein Element der Grundgesamtheit höchstens einmal oder aber auch mehrfach in einer Stichprobe vertreten sein kann. Bei einer Stichprobe ohne Zurücklegen kann eine Untersuchungseinheit nur ein einziges Mal in der Stichprobe vorhanden sein, bei einer Stichprobe mit Zurücklegen mehrmals.

Beispiel 3.9 *Wenn in einem Altenzentrum mit 350 Patient(inn)en, die damit die Grundgesamtheit bilden, insgesamt 25 Patient(inn)en zufällig ausgewählt um im Rahmen einer Befragung die Patientenzufriedenheit zu ermitteln, handelt es sich um eine Stichprobe ohne Zurücklegen.*

Beispiel 3.10 *Werden in einem Krankenhaus n=25 Patient(inn)en einer Station täglich nach ihrem Getränkewunsch zum Frühstück befragt, so handelt es sich um Stichprobe mit Zurücklegen, da die einzelnen Patient(inn)en mehrfach befragt werden.*

Man spricht von einer einfachen Zufallsauswahl oder einfachen Zufallsstichprobe, falls alle n-elementigen Teilmengen der Grundgesamtheit die gleiche Wahrscheinlichkeit besit-